MAKING THE NEXT STEP:
USING EVALUABILITY ASSESSMENT
TO IMPROVE CORRECTIONAL
PROGRAMMING

BETSY MATTHEWS
Eastern Kentucky University

DANA JONES HUBBARD
EDWARD LATESSA
University of Cincinnati

The research on what works in correctional interventions provides a powerful agenda for correctional programming. Evaluability assessment is a tool that can be used to help put this research into practice by providing a measure of program quality. This article describes Gendreau and Andrews's Correctional Program Assessment Inventory (CPAI) as one example of an evaluability assessment tool that is designed to ascertain how well programs are meeting certain principles of effective intervention. It also reports the results of CPAIs conducted on 86 treatment programs. The results indicate that these programs are not adequately incorporating the principles into their correctional programming. Common shortcomings are discussed, and potential resolutions are offered.

Thanks to a growing body of literature on what works in correctional intervention, corrections professionals have for the first time a blueprint for improving correctional programming. Unfortunately, lengthy planning and development processes are uncommon luxuries in the correctional arena. Despite the best of intentions, the pursuit of long-term endeavors is often sacrificed to resolving the more immediate problems of staff turnover, probation violations, and institutional misconduct. Even if programs were afforded the ability to devote time to improving their program, the prospect of developing a program that works can be overwhelming. What is needed is a tool for structuring the inquiry—a tool that can answer the following three basic questions: Where are we now? Where do we need to go? How can we get there?

One tool capable of answering these questions and facilitating program improvements in corrections has been referred to in evaluation literature as an *evaluability assessment* (Prosavac & Carey, 1992; Rutman, 1980; Van Voorhis & Brown, 1996). This article first summarizes the principles of effective intervention as empirically based standards that can serve as benchmarks for program development. Second, it describes an evaluability assessment tool that is designed to provide a standardized quantifiable measure of program quality. Third, it reports the results of 86 assessments of correctional programs across the United States. Fourth, it reports common problems in correctional programming that were identified through the assessments and offers recommendations for improvements. Finally, the article concludes with a discussion regarding the potential benefits of an evaluability assessment tool that provides specific and timely feedback to correctional programs that are interested in taking the next step toward developing programs capable of reducing recidivism.

PRINCIPLES OF EFFECTIVE INTERVENTION

Over the past two decades, numerous literature reviews and meta-analyses have been conducted to examine the effectiveness of various correctional interventions. Despite the differences in samples and statistical techniques, the conclusions drawn by these authors are strikingly similar: Programs that reduce recidivism possess several common characteristics (Andrews, Zinger, et al., 1990; Gendreau & Andrews, 1990; Izzo & Ross, 1990; Lipsey & Wilson, 1998). The most effective programs were conducted in the community (Izzo & Ross, 1990; Lipsey & Wilson, 1998; Palmer, 1974; Whitehead & Lab, 1989), included multimodal programming (Clements, 1988; Lipsey, 1992; Lipsey & Wilson, 1998; Palmer, 1992, 1996), and involved the family in the offender's treatment (Clements, 1988; Gendreau & Ross, 1987; Palmer, 1996).

Other more specific characteristics have been identified. They are referred to as *the principles of effective intervention* (Andrews, Bonta, & Hoge, 1990; Gendreau, 1996; Gendreau & Andrews, 1990). They include the following:

1. Effective interventions are behavioral in nature. A well-designed behavioral program combines a system of reinforcement with modeling by the treatment provider to teach and motivate offenders to perform prosocial behaviors. In addition, problem solving and self-instructional training may be used to change the offenders' cognitions, attitudes, and values that maintain antisocial behavior.

2. Levels of service should be matched to the risk level of the offender. Intensive services are necessary for a significant reduction of recidivism among high-risk offenders, but when applied to low-risk offenders, intensive services produce a minimal or negative effect.

3. Offenders should be matched to services designed to improve their specific criminogenic needs such as antisocial attitudes, substance abuse, family communication, and peer association. Improvements in these areas will contribute to a reduced likelihood of recidivism.

4. Treatment approaches and service providers are matched to the learning style or personality of the offender. For example, high-anxiety offenders do not generally respond well to confrontation (Warren, 1983), and offenders with below-average intellectual abilities do not respond to cognitive skills programs as well as offenders with above average or high intellectual abilities (Fabiano, Porporino, & Robinson, 1991).

5. Services for high-risk offenders should be intensive, occupying 40% to 70% of the offenders' time over a 3- to 9-month period.

6. The program is highly structured, and contingencies are enforced in a firm but fair way: Staff members design, maintain, and enforce contingencies; internal controls are established to detect possible antisocial activities; and program activities disrupt the criminal network and prevent negative peers from taking over the program.

7. Staff members relate to offenders in interpersonally sensitive and constructive ways and are trained and supervised appropriately.

8. Staff members monitor offender change on intermediate targets of treatment.

9. Relapse prevention and aftercare services are employed in the community to monitor and anticipate problem situations and to train offenders to rehearse alternative behaviors.

10. Family members or significant others are trained how to assist clients during problem situations.

11. High levels of advocacy and brokerage occur if community services are appropriate.

Meta-analyses of correctional interventions have found that programs that meet these principles are achieving on average a recidivism reduction of 30% to 50% (Andrews, Zinger, et al., 1990). This research on what works in correctional interventions provides a powerful agenda for correctional programming. The next section of this article describes a valuable tool that can be used to help programs put these principles into practice.

EVALUABILITY ASSESSMENT

According to Van Voorhis and Brown (1996), a major impediment to conducting useful evaluations is flawed program design and implementation.

Poor outcomes are often assumed to be program failures, when in actuality, the program was never implemented as designed (Van Voorhis, Cullen, & Applegate, 1995). Another problem noted by Van Voorhis et al. (1995) is that it is common to see reports on outcome data with no clear indication of what the program did to achieve these results. This approach to evaluation provides a limited basis for program improvements, and it makes the replication of effective programs nearly impossible.

One way to avoid these problems is to conduct an evaluability assessment (Prosavac & Carey, 1992; Rutman, 1980). Evaluability assessments are designed to ascertain if a program has a sound theoretical basis and a well-designed treatment protocol, if the program has been implemented as designed, and if the program is suitable for further inquiry such as an outcome evaluation (Prosavac & Carey, 1992; Van Voorhis et al., 1995). Evaluability assessment comes in many forms. Van Voorhis and Brown (1996) suggested the following four key steps for conducting an evaluability assessment: (a) identifying the purpose and scope of the assessment; (b) developing a program template that describes the goals and objectives of the program, the theory underlying the program, and the intended treatment protocol; (c) validating the program design through interviews and focus groups with program staff members and stakeholders and through observations of program activities; and (d) preparing a report that details the assessment findings and provides appropriate recommendations for future evaluation or program improvements.

The Correctional Program Assessment Inventory (CPAI) (Gendreau & Andrews, 1994)[1] is one example of an evaluability assessment tool that can be used to identify programmatic areas requiring modification. In the case of the CPAI, the principles of effective intervention outlined earlier serve as the program template, and the goal of the assessment is to ascertain the extent to which correctional programs have incorporated these principles. The CPAI provides a standardized, objective way for assessing the quality of correctional programs against empirically based standards. According to Lipsey and Wilson (1998), effective intervention requires a good match between program concept, host organization, and the clientele targeted. The CPAI provides a method for determining if this match exists.

With this type of assessment, agencies can be made aware of inconsistencies in program design and practice before investing in an expensive outcome evaluation. It also facilitates program replication should later outcome evaluations be favorable. The remainder of this article describes the methodology and results of CPAIs conducted on 86 correctional programs and demonstrates the utility of such an assessment.

TABLE 1: Program Type (*N* = 86)

Characteristic	Frequency	Percentage
Service provided		
General services	49	56.9
Substance abuse treatment	34	39.5
Education	2	2.3
Domestic violence	1	1.1
Population		
Juvenile	30	34.5
Adult	56	65.5
Setting		
Residential	46	53.5
Nonresidential	38	44.2
Both	1	1.1
Missing	1	1.1

METHOD

SAMPLE

CPAIs were conducted on 86 programs from January 1996 through September 1998. The sample included both residential and nonresidential programs that provided a variety of services to juvenile and adult offenders. Programs included in this study were not selected randomly; rather, they were self-selected (they requested to be assessed) or chosen by the funding source to be evaluated. The 86 programs included private, nonprofit, and public programs located in 25 states.

Table 1 reports the frequencies and percentages of the services provided, the populations served, and the settings of the programs that were assessed. The majority of the programs (56.9%) provided a variety of services, whereas 39.5% of the programs provided only substance abuse treatment. Two of the programs were limited to the provision of educational services, and one of the programs provided domestic violence treatment. Most of the programs served adult populations (65.5%), and a little more than half of the programs (53.5%) were conducted in residential settings.

MEASURE

The CPAI (Gendreau & Andrews, 1994) was used to ascertain how closely the programs met known principles of effective correctional treatment. There are six primary sections of the CPAI, and they are described as follows:

1. Program implementation: This section focuses on the qualifications and involvement of the program director, the extent to which the treatment literature was considered in the program design, and whether the program is consistent with existing values in the community, meets a local need, and is perceived to be cost-effective.
2. Client preservice assessment: This section examines the program's offender selection and assessment processes to ascertain the extent to which clients are appropriate for the services provided. It also addresses the methods for assessing risk, need, and responsivity factors.
3. Characteristics of the program: This section examines whether the program is targeting criminogenic attitudes and behaviors, the specific treatment modalities employed, the use of rewards and punishments, and the methods used to prepare the offender for release from the program.
4. Characteristics and practices of the staff: This section concerns the qualifications, experience, stability, training, and involvement of the program staff members.
5. Evaluation: This section centers on the types of feedback, assessment, and evaluations used to monitor how well the program is functioning.
6. Miscellaneous: This final section of the CPAI includes miscellaneous items pertaining to the program such as ethical guidelines and levels of funding and community support.

Each section consists of 6 to 22 items for a total of 66 items that are designed to operationalize the principles of effective intervention. The number of items in each section (see Table 2) represents the weight given to that particular section relative to the other sections of the instrument. Each of these items is scored as 1 or 0. To receive a 1, programs must demonstrate that they meet the specified criteria (e.g., the director is involved in some aspect of direct service delivery to clients and client risk of recidivism is assessed through a standardized, quantifiable measure). Based on the percentage of points earned, each section is scored as either *very satisfactory* (70% to 100%), *satisfactory* (60% to 69%), *needs improvement* (50% to 59%), or *unsatisfactory* (less than 50%).[2] The scores from all six areas are totaled, and the same scale is used for the overall assessment score. Some items may be considered not applicable, in which case they are not included in the scoring.

DATA COLLECTION

Teams of researchers visited the 86 program sites. The primary data source was structured interviews with program staff members at each of the sites. Attempts were made to gain information from a representative cross-section of program staff members. Interviews were conducted separately to encourage open discussion and to provide a method for cross-referencing

TABLE 2: Correctional Program Assessment Inventory (CPAI) Results—Category and Mean Scores ($N = 86$)

CPAI Area	Very Satisfactory (70% to 100%)		Satisfactory (60% to 69%)		Needs Improvement (50% to 59%)		Unsatisfactory (< 50%)		M
	%	n	%	n	%	n	%	n	
Program implementation (11 items)	69.8	60	11.6	10	14	12	4.7	4	72.9
Client preservice assessment (11 items)	27.9	24	1.2	1	16.3	14	54.7	47	47.9
Program characteristics (22 items)	7.0	6	12.8	11	25.6	22	54.7	47	43.9
Staff member characteristics (8 items)	34.9	30	17.4	15	25.6	22	22.1	19	60.1
Evaluation (8 items)	14.0	12	2.3	2	10.5	9	73.3	63	38.3
Other (6 items)	74.4	64	17.4	15	7.0	6	1.2	1	83.9
Overall (66 items)	10.6	9	27.1	23	28.2	24	34.1	29	54.8

responses to determine the extent to which consensus exists regarding the purpose and design of the program. Other sources of information included the examination of program documentation, the review of representative case files, and some observation of program activities.

RELIABILITY

Reliability is an important issue in any standardized assessment. Because of the qualitative nature of the data used to score the CPAI, reliability is an important issue. Several steps were taken to increase the reliability of the CPAI. First, to conduct the CPAI, researchers must have demonstrated a strong knowledge of the literature on correctional rehabilitation. Second, researchers participated in a comprehensive training session on how to use the instrument. This training focused on defining the criteria that support each principle of effective intervention, interviewing and observation techniques for collecting the necessary data, and scoring the instrument. Third, each researcher was provided with a scoring guide that listed the criteria that a program must meet to earn a score of 1. Finally, at least two researchers conducted each assessment. Following the scheduled interviews and observations, the researchers compared notes and scores. When inconsistencies were identified, more data were collected to clarify whether the program met the specified criteria.

ANALYSIS

All of the CPAI scores were entered into a project database. Descriptive statistics were used to summarize the results of the assessments. *T* tests were used to identify any significant differences in the average scores between juvenile and adult and residential and nonresidential programs in each area of the CPAI. Individual items and sections were also examined to identify common problem areas among the programs assessed.

RESULTS AND DISCUSSION

CPAI SCORES

Table 2 presents the number and percentage of programs scoring either *very satisfactory, satisfactory, needs improvement*, or *unsatisfactory* on each section of the CPAI and overall. It also reports the mean scores (i.e., the average percentage of points earned) in each area of the CPAI and the mean overall score.

The two strongest areas in the CPAI are other and program implementation. Mean scores of 83.91 and 72.99, respectively, suggest that programs generally meet the principles of effective intervention in these areas. In the area of other, 74% of the programs scored in the very satisfactory range of the scale; only 1.2% of the programs fell in the unsatisfactory range of the scale. This suggests that in general, the programs assessed maintained comprehensive client information in confidential files, had a code of ethics to guide program practices, and demonstrated stability in programming, funding, and community support. In the area of program implementation, 69.8% of the programs fell in the very satisfactory range of the scale; only 4.7% of the programs fell in the unsatisfactory range of the scale. This suggests that in general, the programs assessed had an experienced and involved program director, were developed to meet a pressing local need, and were perceived as being cost-effective and sustainable.

The next strongest area of the CPAI was staff characteristics. Thirty-five percent of the programs assessed scored in the very satisfactory range of the scale, and 22.1% of the programs scored in the unsatisfactory range of the scale. A mean score of 60.12 indicates that most of the programs assessed hired qualified staff members, involved staff members in decision making, and annually evaluated staff members on skills related to service delivery.

The three weakest areas of the CPAI were client preservice assessment, program characteristics, and evaluation. The majority of the programs assessed (54.7% to 74.3%) scored in the unsatisfactory range of the scale in each of these areas. Mean scores ranging from 38.39 in the evaluation section to 47.93 in the assessment section suggest that programs were not committed to evaluating program practices, that the treatment modalities and associated practices lacked a strong theoretical basis, and that programs had not implemented standardized, objective methods for assessing the risk, need, and responsivity factors of clients.

The overall CPAI scores indicate that 34.1% of the programs scored in the unsatisfactory range of the scale. A mean overall score of 54.87 on the CPAI suggests that on average, the programs that were assessed suffered from a general lack of program integrity.

DIFFERENCES IN JUVENILE AND ADULT PROGRAMS

t tests were conducted to examine differences in the CPAI scores between adult and juvenile programs. Table 3 reveals that for the most part, adult and juvenile programs shared the same general strengths and problems. In two of the areas, however, the mean scores for the adult programs were significantly higher than the mean scores for the juvenile programs. First, in the area of

TABLE 3:   *t* Tests on the Differences Between Adult and Juvenile Programs

Correctional Program Assessment Inventory Area	Adult Programs Mean (n = 56)	Juvenile Programs Mean (n = 30)	t-Value	df
Program implementation	75.21	68.85	2.065*	84
Client preservice assessment	51.92	40.53	2.407*	84
Program characteristics	43.26	45.14	−0.476	84
Staff member characteristics	61.14	58.20	0.637	84
Evaluation	39.84	35.69	0.760	84
Other	84.81	82.21	0.726	84
Overall	55.95	52.85	1.236	84

program implementation, the mean score for the adult programs was 75.21 as compared to a mean score of 68.5 for the juvenile programs ($p = .042$). This difference was not attributable to any one factor. Second, in the area of client preservice assessment, the mean score for the adult programs was 51.92 as compared to a mean score of 40.53 for the juvenile programs ($p = .018$). This difference can be attributed to the fact that many juvenile programs have only recently begun to implement standardized, objective methods of risk and need assessment.

DIFFERENCES IN RESIDENTIAL AND NONRESIDENTIAL PROGRAMS

Table 4 presents the findings of *t* tests that were conducted to examine differences between residential and nonresidential programs. The only significant difference in the two samples was in the area of program characteristics. Residential programs had a significantly higher mean score in this area (47.90) as compared to nonresidential programs (39.05; $p = .017$). The difference in the scores on program characteristics can be attributed to the more controlled setting of residential programs. This setting allows for closer monitoring of clients and promotes consistency in the administration of rewards and punishments, both of which are essential ingredients for effective intervention.

COMMON SHORTCOMINGS, IMPLICATIONS, AND POTENTIAL RESOLUTIONS

Table 5 identifies the specific items in the CPAI on which programs generally score poorly. In the area of program implementation, the majority of the programs that were assessed scored poorly on two items; 61.6% failed to conduct a review of pertinent treatment literature, and 62.8% failed to con-

TABLE 4: *t* Tests on Differences Between Nonresidential and Residential Programs (*N* = 86)

Correctional Program Assessment Inventory Area	Nonresidential Mean (n = 38)	Residential Mean (n = 46)	t-Value	df
Program implementation	73.69	72.10	0.528	82
Client preservice assessment	49.06	47.26	2.407	82
Program characteristics	39.05	47.90	−2.443*	82
Staff member characteristics	62.85	57.76	1.151	82
Evaluation	35.95	40.85	−0.919	82
Other	83.16	85.28	−0.621	82
Overall	53.58	56.10	−1.066	82

duct a pilot period prior to implementation. Without a review of the treatment literature, program models are not likely to reflect theories and practices that have been found to be effective in reducing recidivism. Furthermore, a pilot period is needed to sort out program logistics and content prior to the formal implementation of the program. Changes can be difficult to initiate once a program is formally implemented. Existing programs are advised to review the literature pertinent to program modifications or additions and to pilot new program components for at least 1 month.

The programs scored poorly on several items in the area of client preservice assessment. Although most programs assess clients' risks and needs associated with recidivism, they generally do so with a clinical psychosocial assessment that is based on subjective, professional judgment. Empirical evidence suggests that actuarial methods of assessment that are based on objective and standardized factors are more accurate at predicting the risk of recidivism than are clinical methods (Gottfredson, 1987; Jones, 1996). Only 33.7% of the programs, however, used an objective method of risk assessment, and only 41.9% used an objective method of need assessment. Without an objective method for risk and need assessment, these programs are left without a summary score that defines the clients' level of risk and need and that can be used in case classification. This impedes a program's ability to allocate treatment resources in an equitable manner as well as its ability to assign clients to the intensity and type of treatment that best matches their risks and needs. It is recommended that programs develop actuarial methods of assessment to identify the general risks and needs of clients. Examples of such instruments include the Wisconsin Risk/Need Instrument (Baird, Heinz, & Bemus, 1979), the Level of Services Inventory-Revised (LSI) (Andrews & Bonta, 1995), or the Correctional Offender Management Profiles for Alternative Sentences (Northpointe Institute for Public Management, 1997). Each of these instruments includes factors that are known corre-

TABLE 5:   Correctional Program Assessment Inventory (CPAI) Items (or Principles of Effective Intervention) That Programs Are Not Meeting ($N = 86$)

	No		Yes	
	%	n	%	n
Program implementation				
Literature review	61.6	53	34.9	30
Pilot period	62.8	54	18.6	16
Client preservice assessment				
Risk methods	64.0	55	33.7	29
Need methods	57.0	49	41.9	36
Responsivity factors	74.4	64	26.6	22
Responsivity methods	80.2	69	17.4	15
Risk level defined	66.3	57	32.6	28
Need level defined	64.0	55	36.0	31
Responsivity defined	83.7	72	16.3	14
Program characteristics				
Matching risk level to service	82.6	71	16.3	14
Matching treatment to client	88.4	76	11.6	10
Matching staff members to treatment	75.6	65	24.4	21
Matching staff members to client	89.5	77	10.5	9
Ratio favors rewards	84.9	73	15.1	13
Theory of punishment	84.9	73	12.8	11
Appropriate punishing stimuli	63.5	54	29.4	25
Procedure for punishment	70.6	60	22.4	19
Negative effects	80.0	68	15.3	13
Significant others	69.8	60	27.9	24
Booster/aftercare provided	76.4	64	23.3	20
Staff member characteristics				
Training for 3 to 6 months	87.2	75	9.3	8
Evaluation				
Of clients on target areas	70.9	61	29.1	25
Clients followed	54.7	47	37.2	32
Outcome study (methodological)	70.9	61	9.3	8

NOTE: Percentages may not add to 100% because some programs were scored not applicable in some of the aforementioned CPAI items.

lates of crime and dynamic factors that are amenable to change, and each instrument provides a summary score for use in case classification. Several researchers have cautioned agencies against adopting existing instruments for use with their offender population and warn that accurate prediction is dependent on an instrument that is validated and normed on the population on which it is to be used (Clear, 1988; Gottfredson, 1987; Jones, 1996). If specific areas of need are identified through this initial assessment (e.g., substance abuse, aggression, and antisocial values), other more in-depth assessments could then be conducted to determine the extent of the problem.

Only 26.6% of the programs assessed responsivity factors, and those programs that did assess responsivity factors did not use a standardized, objec-

tive method of assessment. Responsivity factors are traits of the client that may interfere with treatment effectiveness. For example, persons with low intelligence do not perform well in cognitive skills programs (Fabiano et al., 1991), high-anxiety offenders do not respond well to counselors or treatment environments that are highly confrontational (Warren, 1983), and persons with a low conceptual level need a treatment strategy that provides a high level of structure (Brill, 1978; Reitsma-Street & Leschied, 1988). Each of these factors could be assessed at intake. In addition, research has identified subgroups of offenders with varying personality traits that influence how they respond to treatment (Van Voorhis, 1994). Two instruments that can be used to identify these subgroups of offenders include the Client Management Classification Assessment Instrument (Lerner, Arling, & Baird, 1986) and the Jesness Inventory (Jesness, 1988). Both of these instruments offer intervention strategies for each of the subgroups of offenders.

The problems in the area of client preservice assessment carry over to the area of program characteristics. Without standardized, quantifiable methods of assessment, programs are limited in their ability to match clients to appropriate services. Even when programs have a comprehensive assessment process, however, little treatment matching occurs; programs tend to use a one-size-fits-all approach. One of the most important principles of classification is the risk principle, which states that the level of service should be matched to the risk level of the offender (Andrews, Bonta, & Hoge, 1990). Only 16.3% of the programs varied the intensity and duration of the program according to the clients' risk levels. This common shortcoming has serious ramifications. The application of intensive services and controls to low-risk offenders can actually be harmful; it interferes with the generally prosocial lifestyles of these offenders and in some cases increases their risk of recidivism (Andrews, Bonta, et al., 1990; Clear & Hardyman, 1990). Furthermore, according to Gendreau (1996), intensive services that occupy at least 40% of an offender's time are needed to produce significant reductions in the recidivism of high-risk offenders.

Another important principle of classification is the responsivity principle, which states that offenders should be matched to treatment approaches and to treatment staff members that accommodate their learning styles and personality traits (Andrews, Bonta, et al., 1990). Only 11.6% of the programs matched clients to treatment approaches based on responsivity factors, and only 10.5% of the programs matched clients to treatment staff members. This failure to match treatment strategies with the personal characteristics of offenders can lead to an increase in program failures. Many programs appear ineffective when it may be that inappropriate cases are masking the success that these programs are having with the clients who are appropriate for the ser-

vices provided (Van Voorhis et al., 1995). The advantages of incorporating the responsivity principle are twofold: First, agencies benefit from increased program success, and second, clients benefit from the provision of services that are well suited to their unique constellation of needs and personality characteristics.

Another common shortcoming in the area of program characteristics concerns a program's failure to systematically match staff members with the types of treatment that they provide. This matching occurred in only 24.4% of the programs assessed, with most of the agencies appearing to be more interested in having staff members be jacks-of-all-trades. Although there are obvious advantages to this approach (e.g., staff coverage), the disadvantages are more compelling. First, staff members are not given the opportunity to develop a particular area of expertise. Second, and more importantly, the failure to match staff members to services based on specific skills and interests can threaten program integrity and interfere with client treatment. The optimal time to implement this matching principle is during the recruitment and hiring stage. It can be effectively implemented, however, through ongoing staff member evaluation and professional development that is designed to ensure that staff members have the requisite skills and motivation for delivering the designated services.

One of the most problematic areas within the program characteristics portion of the CPAI revolves around the use of rewards and punishments. As suggested by the literature on behavioral and social learning theories, positive reinforcement is a powerful mechanism of behavioral change (Akers, 1985; Spiegler & Guevremont, 1993). Although most programs have some appropriate rewards in place, their use is somewhat sporadic. It is recommended that the ratio of rewards to punishers be at least 4:1 (Gendreau, 1996); only 15.1% of the agencies that were assessed met this criterion. It is recommended that a range of possible rewards be identified and that program policies be developed to encourage their systematic application. In addition, given the context of correctional programming, staff members may require training on the importance of positive reinforcement to overcome their predominant focus on negative behavior and their dependence on negative sanctions as a means to control behavior.

Despite the overreliance on punishment, it appears that correctional programs lack theoretical insight on why they use punishment and practical knowledge about how to punish effectively. Only 12.8% of the programs had a theoretically relevant rationale for their use of punishment. Rather than using punishment to extinguish or suppress negative behavior over the long term, punishment was more often used as a means of short-term, in-program control of behavior. Moreover, only 22.4% of the programs followed the principles of effective punishment. As outlined by Gendreau (1996), six cri-

teria must be met for punishment to be effective. First, punishment must be administered at the earliest point in the deviant behavior. Second, punishment must be administered after each occurrence of the deviant behavior. Third, the punishment should be administered at maximum intensity. Fourth, the punishment should vary with repeated negative behavior. Fifth, escape from the punishing stimuli must be impossible. Sixth, in addition to the punishment, the client must be taught a prosocial alternative to the deviant behavior. The CPAIs also revealed insufficient follow-up to punishment. Only 15.3% of the programs monitored clients to look for any unintended negative effects of punishment such as aggression or withdrawal. It is strongly recommended that correctional agencies interested in using punishment as a mechanism of behavioral change draw on the psychological literature (see Gendreau, 1996; Spiegler & Guevremont, 1993) to develop training and policies that encourage the proper administration of punishment.

Two other important components that are difficult to implement in correctional intervention programs are the involvement of families or significant others in the clients' treatment and the provision of aftercare services. Only 27.9% of the programs engaged families in treatment, and when they did offer services to families, they experienced very low participation rates. Juvenile programs appear to have more success in this area. Some programs have experienced success in getting families to participate by making it a condition of the client's successful release from the program. Other programs have had success in getting families involved by providing transportation to the program.

Only 25.6% of the programs assessed provided aftercare or booster sessions. A common finding was that although a residential program may have prepared aftercare plans for clients, the plans were not implemented after the clients' release to probation or parole supervision. The consistent provision of aftercare services requires collaboration among all components of the criminal justice system. Given that meta-analyses of correctional treatment studies have found that the provision of aftercare services is associated with higher reductions in recidivism (Gendreau & Andrews, 1990; Lipsey & Wilson, 1998), this area should be given high priority in program development.

In the area of staff member characteristics, programs consistently had trouble meeting the criterion for staff member training. The CPAI criterion states that new staff members should receive 3 to 6 months of formal training on the type of intervention being used in the program. Only 9.3% of the programs met this criterion. All too often, the only training staff members received was on-the-job training that entailed shadowing an experienced worker and reading program materials. It is recommended that this type of

on-the-job training be accompanied by formal training on the theories of the intervention (e.g., behavioral-cognitive and social learning theories).

There were several problem areas in the evaluation component of the CPAI. First, only 29.1% of the programs evaluate clients on their improvements in target areas. It is recommended that standardized and quantifiable instruments be used as preintervention and postintervention measures to assess changes in key target areas. For example, one of the programs used the LSI at intake and discharge to assess change in client risk. This procedure offered the following two benefits: It allowed them to monitor individual client progress, and it provided them with data on which to assess specific program elements. This particular program discovered that no significant improvements were occurring in the area of peer relationships. Based on these data, they modified their treatment curriculum to include a component that addressed the issue of peer pressure. Using preintervention and post-intervention measures is one of the most efficient and reliable methods of program evaluation.

Second, only 37.2% of the programs collected follow-up data on their clients for at least 6 months after release from the program. Finally, only 9.3% of the programs have conducted outcome studies. The second and third problems with the evaluation component are not so easily resolved. Both are time-consuming and costly endeavors.

SUMMARY AND CONCLUSIONS

The overall CPAI scores indicate that 34.1% of the programs scored in the unsatisfactory range of the scale. A mean overall score of 54.87 on the CPAI suggests that on average, the programs that were assessed suffered from a general lack of program integrity. The good news is that the requisite organizational framework appeared to be in place: Program directors and staff members were well qualified and integral to program development, the level of funding was adequate and sustainable, and there was internal and external support for the programs.

On the other hand, the majority of programs we assessed do not adequately assess offender risk, need, or responsivity factors; do not utilize effective treatment models; do not use behavioral strategies; do not adequately train staff members; and do not evaluate themselves or the performance of the offenders they serve. These findings mirror other studies of program integrity (Gendreau & Goggin, 2000; Latessa & Holsinger, 1998).

There are a number of limitations to the CPAI that should be noted. First, the instrument is more applicable for a self-contained program and is more

difficult to apply on certain programs. For example, when examining an agency or program that uses several external treatment providers, it is necessary to examine each component, which in turn may affect the scoring of the instrument. Second, the instrument is based on the ideal type of program. It is unlikely that any correctional program can meet all of the criteria set forth in the CPAI. Third, reliability can be a problem because the process involves making some decisions concerning the information obtained in the interviews. As mentioned previously, steps are taken to increase reliability in scoring; however, when there are disagreements in scoring, a final decision is ultimately made. Fourth, the assessment is present oriented; the score is based on the program at the time of the assessment. Accordingly, the assessment might give a program credit for recent changes although the improvements in the program may have just been made. Fifth, the process does not take into account system issues. For example, a program's effectiveness can be undermined because of a lack of understanding or support from others within the system. Often, programs have little control over these factors.

Despite these limitations, there are a number of advantages to this process. First, the instrument is based on empirically derived principles that are applicable to a wide range of correctional programs. Second, the instrument provides a measure of program integrity—the degree to which a program meets the principles of effective intervention—and a measure of program quality—the extent to which a program consistently delivers services and interventions. Third, the process identifies both strengths and weaknesses of a program. In other words, it identifies those areas of a program that are consistent with the principles of effective intervention as well as areas that are not. Fourth, a product of the CPAI is a detailed written report that provides recommendations for program improvement. These recommendations offer specific ways that agencies can meet the principles of effective intervention by improving or adding components and practices. Fifth, because programs are assessed on the same criteria and the process is quantifiable, the process can be used for benchmarking. Program scores can be compared over time as well as across programs. Finally, with baseline information on the quality of their programs, many of these agencies have the capacity to move forward and implement practices that reflect the principles of effective correctional programming.

NOTES

1. The Correctional Program Assessment Inventory (CPAI) is a copyrighted instrument. Its use requires training and the permission of Paul Gendreau and Don Andrews.

2. The cutoff scores were arbitrarily derived; however, because the CPAI is based on an ideal type program, it is unlikely that any program will ever achieve a perfect score. In addition, a recent validity study of the instrument demonstrated a high degree of correlation between CPAI scores and offender outcome. The relationship was generally linear and further supports the established categories (see Holsinger, 1999).

REFERENCES

Akers, R. L. (1985). *Deviant behavior: A social learning approach* (3rd ed.). Belmont, CA: Wadsworth.

Andrews, D. A., & Bonta, J. (1995). *The Level of Supervision Inventory-Revised.* Toronto, Canada: Multi-Health Systems.

Andrews, D. A., Bonta, J., & Hoge, R. D. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior, 17,* 19-52.

Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. T. (1990). Does correctional treatment work? Clinically relevant and psychologically informed meta-analysis. *Criminology, 28,* 369-404.

Baird, S. C., Heinz, R., & Bemus, B. (1979). *The Wisconsin case classification/staff deployment project* (Project Report No. 14). Madison, WI: Department of Health and Social Services, Division of Corrections.

Brill, R. (1978). Implications of the conceptual level matching model for treatment of delinquents. *Journal of Research in Crime and Delinquency, 15,* 229-246.

Clear, T. (1988). Statistical prediction in corrections. *Research in Corrections, 1,* 1-39.

Clear, T., & Hardyman, P. (1990). The new intensive supervision movement. *Crime and Delinquency, 36,* 42-60.

Clements, C. (1988). Delinquency prevention and treatment: A community-centered perspective. *Criminal Justice and Behavior, 15,* 286-305.

Fabiano, E., Porporino, F., & Robinson, D. (1991). Canada's cognitive skills program corrects offenders' faulty thinking. *Corrections Today, 53,* 102-108.

Gendreau, P. (1996). The principles of effective intervention with offenders. In A. T. Harland (Ed.), *Choosing correctional options that work* (pp. 117-130). Thousand Oaks, CA: Sage.

Gendreau, P., & Andrews, D. A. (1990). Tertiary prevention: What the meta-analyses of the offender treatment literature tell us about what works. *Canadian Journal of Criminology, 32,* 173-184.

Gendreau, P., & Andrews, D. A. (1994). *Correctional Program Assessment Inventory* (4th ed.). St. John, New Brunswick, Canada: University of New Brunswick.

Gendreau, P., & Goggin, C. (2000). Correctional treatment: Accomplishments and realities. In P. Van Voorhis, M. Braswell, & D. Lester (Eds.), *Correctional counseling and rehabilitation* (pp. 289-298). Cincinnati, OH: Anderson.

Gendreau, P., & Ross, R. (1987). Revivification of rehabilitation: Evidence from the 1980s. *Justice Quarterly, 4,* 349-407.

Gottfredson, D. (1987). Prediction and classification. In D. Gottfredson & M. Tonry (Eds.), *Prediction and classification: Criminal justice decision making* (pp. 1-20). Chicago: University of Chicago Press.

Holsinger, A. (1999). *Opening the black box: Assessing the relationship between program integrity and recidivism.* Unpublished doctoral dissertation, University of Cincinnati.

Izzo, R., & Ross, R. (1990). Meta-analysis of rehabilitation programs for juvenile delinquents: A brief report. *Criminal Justice and Behavior, 17,* 134-142.

Jesness, C. F. (1988). Jesness Inventory Classification System. *Criminal Justice and Behavior, 15*, 78-91.

Jones, P. (1996). Risk prediction in criminal justice. In A. T. Harland (Ed.), *Choosing correctional options that work* (pp. 33-68). Thousand Oaks, CA: Sage.

Latessa, E. J., & Holsinger, A. (1998). The importance of evaluating correctional programs: Assessing outcome and quality. *Corrections Management Quarterly, 2*(4), 22-29.

Lerner, K., Arling, G., & Baird, S. C. (1986). Client management classification strategies for case supervision. *Crime and Delinquency, 32*, 254-271.

Lipsey, M. (1992). Juvenile delinquency treatment: A meta-analytic inquiry into the variability of effects. In T. Cook, H. Cooper, D. Cordray, H. Hartman, L. Hedges, R. Light, T. Louis, & F. Mosteller (Eds.), *Meta-analysis for explanation* (pp. 83-127). New York: Russell Sage.

Lipsey, M., & Wilson, D. (1998). Effective intervention for serious juvenile offenders: A synthesis of research. In R. Loeber & D. P. Farrington (Eds.), *Serious and violent juvenile offenders: Risk factors and successful interventions* (pp. 313-345). Thousand Oaks, CA: Sage.

Palmer, T. (1974). The Youth Authority's community treatment project. *Federal Probation, 38*, 3-14.

Palmer, T. (1992). *The re-emergence of correctional intervention.* Newbury Park, CA: Sage.

Palmer, T. (1996). Programmatic and nonprogrammatic aspects of successful intervention. In A. T. Harland (Ed.), *Choosing correctional options that work* (pp. 131-182). Thousand Oaks, CA: Sage

Prosavac, E. J., & Carey, R. G. (1992). *Program evaluation: Methods and case studies* (4th ed.). Englewood Cliffs, NJ: Prentice Hall.

Reitsma-Street, M., & Leschied, A. W. (1988). The conceptual-level matching model in corrections. *Criminal Justice and Behavior, 15*, 92-108.

Rutman, L. (1980). *Planning useful evaluations: Evaluability assessment.* Beverly Hills, CA: Sage.

Spiegler, M. D., & Guevremont, D. C. (1993). *Contemporary behavior therapy* (2nd ed.). Pacific Grove, CA: Brooks/Cole.

Van Voorhis, P. (1994). *Psychological classification of the adult male prison inmate.* Albany: State University of New York Press.

Van Voorhis, P., & Brown, K. (1996). *Evaluability assessment: A tool for program development in corrections.* Unpublished monograph, National Institute of Corrections, Washington, DC.

Van Voorhis, P., Cullen, F. T., & Applegate, D. (1995). Evaluating interventions with violent offenders. *Federal Probation, 50*, 17-27.

Warren, M. (1983). Application of interpersonal maturity theory to offender populations. In W. Laufer & J. Day (Eds.), *Personality theory, moral development, and criminal behavior* (pp. 23-49). Lexington, MA: Lexington Books.

Whitehead, J., & Lab, S. (1989). A meta-analysis of juvenile correctional treatment. *Journal of Research in Crime and Delinquency, 26*, 276-295.